LIBRARY
RESEARCH REPORTS DIVISION
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943

NPS55-84-019

NAVAL POSTGRADUATE SCHOOL

Monterey, California

TECHNICAL

MODELLING AND RESIDUAL ANALYSIS OF NONLINEAR AUTO-REGRESSIVE TIME SERIES IN EXPONENTIAL VARIABLES

by

P. A. W. Lewis
A. J. Lawrance

August 1984

Approved for public release; distribution unlimited

Prepared for:

Chief of Naval Research
Arlington, Va 22217

FEDDOCS
D 208.14/2:
NPS-55-84-019

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Commodore R. H. Shumaker
Superintendent

David A. Schrady
Provost

This work was supported by the Office of Naval Research under
Grant N00014-84-WR24004

Reproduction of all or part of this report is authorized.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-84-019	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* MODELLING AND RESIDUAL ANALYSIS OF NONLINEAR AUTO-REGRESSIVE TIME SERIES IN EXPONENTIAL VARIABLES.		5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) P. A. W. Lewis A. J. Lawrance		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61152N; RR000-01-10 N0001484WR41001
11. CONTROLLING OFFICE NAME AND ADDRESS Chief of Naval Research Arlington, Va 22217		12. REPORT DATE August 1984 13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

An approach to modelling and residual analysis of nonlinear autoregressive time series in exponential variables is presented; the approach is illustrated by an analysis of a long series of wind velocity data which has first been detrended and then transformed into a stationary series with an exponential marginal distribution. The stationary series is modelled with a newly developed type of second order autoregressive process with random coefficients, called the NEAR(2)model; it has a second order autoregressive correlation structure but is nonlinear because its coefficients are random. The exponential distri-

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

butional assumptions involved in this model highlight a very broad four parameter structure which combines five exponential random variables into a sixth exponential random variable; other applications of this structure are briefly considered. Dependency in the NEAR(2) process not accounted for by standard autocorrelations is explored by developing a residual analysis for time series having autoregressive correlation structure; this involves defining linear uncorrelated residuals which are dependent, and then assessing this higher order dependence by standard time series computations. The application of this residual analysis to the wind velocity data illustrates both the utility and difficulty of nonlinear time series modelling.

28th August 1984

MODELLING AND RESIDUAL ANALYSIS OF NONLINEAR AUTOREGRESSIVE TIME SERIES IN EXPONENTIAL VARIABLES

by

A J Lawrance
Department of Statistics
University of Birmingham
Birmingham, England

P A W Lewis
Department of Operations Research
Naval Postgraduate School
Monterey, California, USA

ABSTRACT

An approach to modelling and residual analysis of nonlinear autoregressive time series in exponential variables is presented; the approach is illustrated by an analysis of a long series of wind velocity data which has first been detrended and then transformed into a stationary series with an exponential marginal distribution. The stationary series is modelled with a newly developed type of second order autoregressive process with random coefficients, called the NEAR(2) model; it has a second order autoregressive correlation structure but is nonlinear because its coefficients are random. The exponential distributional assumptions involved in this model highlight a very broad four parameter structure which combines five exponential random variables into a sixth exponential random variable; other applications of this structure are briefly considered. Dependency in the NEAR(2) process not accounted for by standard autocorrelations is explored by developing a residual analysis for time series having autoregressive correlation structure; this involves defining linear uncorrelated residuals which are dependent, and then assessing this higher order dependence by standard time series computations. The application of this residual analysis to the wind velocity data illustrates both the utility and difficulty of nonlinear time series modelling.

1. INTRODUCTION

There are several aspects of many observed univariate time series which are not satisfactorily accounted for in standard time series analysis: they include nonGaussian marginal distributions, dependence not accounted for by second order moments (autocorrelations) and directionality in the time series. Quite often a Gaussian distribution will be inappropriate because the variable being modelled has a positive and highly skewed distribution, e.g. wind speeds, the service times in a queue, or the daily flows of a river. Many particular such distributions can be envisaged and time series models have been constructed for them. Examples are Gamma distributions (Gaver and Lewis, 1980; Lewis, 1981; McKenzie, 1982, Lawrance, 1982) and mixed exponential distributions (Gaver and Lewis, 1980; Lawrance, 1980a; Lawrance and Lewis, 1982).

However the simplest, most widely used and most analytically tractable of these distribution models is the exponential distribution. Like Gaussian random variables, exponentially distributed random variables enjoy many special properties; also they can be mildly transformed quite easily into distributions which are either more skewed or less skewed than the exponential. The Weibull distribution is an example, being just a power transformation of an exponentially distributed random variable. Thus the approach here, following earlier work (Gaver and Lewis, 1980; Lawrance and Lewis, 1980, 1981) is to regard the exponential variables as canonical and to develop their use in time series modelling.

It should also be noted that time series of (marginally) uniformly distributed random variables can be obtained by exponential transformations of time series in exponentially distributed variables; such uniform processes could then be used to generate time series with other desired marginal distributions.

The work cited previously has concentrated for the most part on first order, nonGaussian autoregressive models, both of the standard type (constant coefficient, additive, linear combinations) and a random coefficient type introduced by the authors. The extension of the models to higher order autoregression is clearly necessary to attain flexibility in modelling correlation and dependency structure of the processes, but these extensions are in no way as immediate as in the standard linear Gaussian case. A simple mixing device can be used (Jacobs and Lewis, 1983) but the range of correlations attained is much narrower than the range attained in the standard linear, second order autoregressive structure. A broader extension, called the EAR(2) model, was obtained in the exponential case by Lawrance and Lewis (1980), but its innovation variable has a zero component which gives runs in the process; this will often be hard to justify.

A major part of the present work consists of obtaining a very broad and rich extension of the NEAR(1) model (Lawrance and Lewis, 1981) to a second order autoregressive process; it includes the EAR(2) model but does not generally have a zero component. This NEAR(2) model was proposed in Lawrance (1980b), later reviewed in Raftery (1981), but the necessary analysis of its innovation structure was not given. Here the innovation random variable for the NEAR(2) process is proved to exist without unnatural boundaries on its (four) parameter region; explicit construction is given for the innovation random variable.

The richness of the four parameter NEAR(2) model, and the fact that an infinite number of cases of the model with identical correlation structure are available, forces consideration of higher order aspects of dependence. The analysis of the higher order aspects of exponential time series is at a fairly early stage and is as follows. First it will be shown that the autocorrelations $\rho(\ell)$, $\ell = 0, \pm1, \pm2, \ldots$ for the NEAR(2) process satisfy the Yule-Walker equations with constants a_1 and a_2 which are functions of the

four parameters of the model. This follows immediately from the fact that X_n is a random ceofficient, linear additive combination of X_{n-1}, X_{n-2} and the innovation random variable ϵ_n. Secondly, it can be shown (Lewis and Lawrance, 1984) that the residuals $X_n - a_1X_{n-1} - a_2X_{n-2}$, which are the usual residuals for second order constant coeficient, linear additive autoregressive processes are *uncorrelated*.

Thus, although the standard analysis of time series stops with uncorrelated residuals, i.e. a flat spectrum for the residuals, such residuals can also be used to good effect to investigate higher order aspects of dependence in the NEAR(2) model. In fact, if the autoregression is not of the standard type (constant coefficient, additive, linear combinations) the (uncorrelated) residual will be *dependent*. One aspect of this is that the squared residuals will have non-zero autocorrelations, and another is that the crosscorrelations of residuals and squared residuals will be non-zero; both sets of correlations are theoretically zero when a standard second order autoregressive model is appropriate.

This residual analysis will be illustrated by some theoretical calculations for the NEAR(2) model and by a brief application to a long series of detrended and transformed wind velocity data.

2. EXPONENTIAL TIME SERIES MODELS

Our aim in this section is to give in outline the ideas leading to the time series model of main concern in this paper, and called NEAR(2), following the earlier terminology NEAR(1) in Lawrance and Lewis (1981). The NEAR(2) model has four parameters, and incorporates and broadens the earlier two parameter EAR(2) model (Lawrance and Lewis, 1980). The NEAR(2) model will be exponential in marginal distribution, have second order autoregressive Markov dependence, and have autocorrelations satisfying second order difference equations of the familiar Yule-Walker type. In addition it will have dependence beyond autocorrelation, and will not be reversible in time. It is not linear in the standard sense, having random coefficient, linear additive autoregressive structure, but neither is it nonlinear in the standard sense of incorporating powers or products of lagged variables. Also, it differs from the random coefficient models considered by Nicholls and Quinn (1982) in that the marginal distribution is specified. The view taken here is that the marginal distribution is the easiest aspect of data to look at and should be the starting point for modelling.

Writing $\{X_n\}$ for the time series variables, and $\{E_n\}$ for an i.i.d. exponential innovation sequence of unit mean, the two parameter NEAR(1) model, as previously defined, is given by

$$X_n = \begin{cases} \beta X_{n-1} & \text{w.p. } \alpha \\ 0 & \text{w.p. } 1-\alpha \end{cases} + \begin{cases} E_n & \text{w.p. } p \\ bE_n & \text{w.p. } 1-p \end{cases}, \qquad (2.1)$$

with $b=(1-\alpha)\beta$ and $p=(1-\beta)/\{1-(1-\alpha)\beta\}$. The parameter region is, in general, $0\leq\alpha,\beta\leq 1$, $\alpha=\beta\neq 1$. The case $\beta=1$, $0\leq\alpha<1$ is rather special, and has been called the TEAR(1) model, and when $\alpha=1$, $0\leq\beta<1$, the earlier EAR(1) model is recovered. *Except for this* EAR(1) *case, the* NEAR(1) *model does not allow zero innovations (Gaver and Lewis, 1980) and so is more statistically acceptable.* The zero innovation in the EAR(1) case implies that $X_n=\beta X_{n-1}$

and thus β can be determined exactly from runs down in the sample path of the process.

In general the i.i.d. innovations in the NEAR(1) process are formed as the probabilistic mixture of two exponentials, and are thus easily simulated.

The NEAR(2) model is a direct generalization of (2.1) and takes the form

$$X_n = \begin{cases} \beta_1 X_{n-1} & \text{w.p. } \alpha_1 \\ \beta_2 X_{n-2} & \text{w.p. } \alpha_2 \\ 0 & \text{w.p. } 1-\alpha_1-\alpha_2 \end{cases} + \epsilon_n \qquad (2.2)$$

with parameter region $\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_1+\alpha_2 \leq 1, 0 \leq \beta_1, \beta_2 \leq 1$; $\{\epsilon_n\}$ is an appropriately chosen innovation sequence. Many special cases can arise when the above restrictions include some of the equalities and, for the purposes of a general development, it is best to regard the inequalities as strict. Given that $\{X_n\}$ is required to have an exponential marginal distribution, the main question concerns whether there is a valid probability distribution for ϵ_n. The Theorem proved in Section 2.3 will show that this is the case, and that the distribution, when the inequalities on α_1, α_2 and β_1, β_2 in the parameter region are strict, takes the form

$$\epsilon_n = \begin{cases} E_n & \text{w.p. } 1-p_2-p_3 \\ b_2 E_n & \text{w.p. } p_2 \\ b_3 E_n & \text{w.p. } p_3 \end{cases}, \qquad (2.3)$$

a probabilistic mixture of three exponentials with parameters given in Section 2.3. To establish this result a fairly detailed analysis of a derived moment generating function is required. This is necessary since a direct moment generating function solution of (2.2) for ϵ_n does not establish that ϵ_n has a proper distribution; all that is shown is that the solution is a possibly-improper mixture of three exponentials.

3. VALIDITY OF THE NEAR(2) MODEL

In this section we prove the following

THEOREM. Let $\{E_n\}$ be an i.i.d. sequence of unit mean exponential random variables. Then if the four parameters $\alpha_1,\alpha_2,\beta_1,\beta_2$ satisfy $\alpha_1>0,\alpha_2>0$, $\alpha_1+\alpha_2<1, 0<\beta_1,\beta_2<1$, the relationship

$$X_n = \left.\begin{cases}\beta_1 X_{n-1} & \text{w.p. } \alpha_1 \\ \beta_2 X_{n-2} & \text{w.p. } \alpha_2 \\ 0 & \text{w.p. } 1-\alpha_1-\alpha_2\end{cases}\right\} + \epsilon_n, \qquad n=0,\pm1,\pm2,\ldots, \tag{3.1}$$

where

$$\epsilon_n = \begin{cases} E_n & \text{w.p. } 1-p_2-p_3 \\ b_2E_n & \text{w.p. } p_2 \\ b_3E_n & \text{w.p. } p_3\end{cases}, \tag{3.2}$$

defines a stationary sequence of (marginally) exponentially distributed random variables with mean one. Here

$$p_2 = \{(\alpha_1\beta_1+\alpha_2\beta_2)b_2 - (\alpha_1+\alpha_2)\beta_1\beta_2\}/\{(b_2-b_3)(1-b_2)\}, \tag{3.3}$$

$$p_3 = \{(\alpha_1+\alpha_2)\beta_1\beta_2 - (\alpha_1\beta_1+\alpha_2\beta_2)b_3\}/\{(b_2-b_3)(1-b_3)\}, \tag{3.4}$$

and

$$0 < b_3 = \{s-(s^2-4r)^{1/2}\}/2 < b_2 = \{s+(s^2-4r)^{1/2}\}/2 < 1, \tag{3.5}$$

where

$$s = (1-\alpha_1)\beta_1 + (1-\alpha_2)\beta_2, \tag{3.6}$$

and

$$r = (1-\alpha_1-\alpha_2)\beta_1\beta_2. \tag{3.7}$$

PROOF. For the NEAR(2) model specified by (3.1)-(3.7), let $\phi_X(t)$ and $\phi_\epsilon(t)$ be the moment generating functions of the $\{X_n\}$ and $\{\epsilon_n\}$ sequences; then if stationarity of the $\{X_n\}$ series is assumed,

$$\phi_X(t) = \phi_\epsilon(t)\{\alpha_1\phi_X(\beta_1 t) + \alpha_2\phi_X(\beta_2 t) + (1-\alpha_1-\alpha_2)\}. \tag{3.8}$$

Assuming an exponential marginal distribution of unit mean for $\{X_n\}$, then the independent distribution of $\{\epsilon_n\}$ has moment generating function, possibly not proper, given by

$$\phi_\epsilon(t) = \frac{(1+\beta_1 t)(1+\beta_2 t)}{(1+t)[(1-\alpha_1-\alpha_2)\beta_1\beta_2 t^2 + \{(1-\alpha_1)\beta_1 + (1-\alpha_2)\beta_2\}t + 1]}. \tag{3.9}$$

It is convenient to establish right away that the quadratic term in the denominator of (3.9) has real distinct and positive roots, b_1 and b_2; this eliminates any subsequent need to invert such a term as a whole. The required condition for real distinct roots is that

$$[(1-\alpha_1)\beta_1 + (1-\alpha_2)\beta_2]^2 - 4(1-\alpha_1-\alpha_2)\beta_1\beta_2$$

be positive: this is so from its equality to the expression

$$[(1-\alpha_1)\beta_1 - (1-\alpha_2)\beta_2]^2 + 4\alpha_1\alpha_2\beta_1\beta_2$$

which is clearly positive; the positivity of the roots b_1 and b_2 is obvious from (3.9) since their product and sum given in (3.11) and (3.12) below are both positve.

With b_1 and b_2 positive, a partial fraction expansion of (3.2) can be written in the suggestive form

$$\Phi_\epsilon(t) = (1-p_2-p_3)\frac{1}{1+t} + p_2\frac{1}{1+b_2t} + p_3\frac{1}{1+b_3t} . \qquad (3.10)$$

Comparisons between (3.9) and (3.10) then show that b_2,b_3 and p_2,p_3 may be obtained in terms of β_1,β_2 and α_1,α_2 by solving the equations

$$b_2 + b_3 = (1-\alpha_1)\beta_1 + (1-\alpha_2)\beta_2, \qquad (3.11)$$

$$b_2b_3 = (1-\alpha_1-\alpha_2)\beta_1\beta_2, \qquad (3.12)$$

$$(1-b_2)p_2 + (1-b_3)p_3 = \alpha_1\beta_1 + \alpha_2\beta_2, \qquad (3.13)$$

$$b_3(1-b_2)p_2 + b_2(1-b_3)p_3 = (\alpha_1+\alpha_2)\beta_1\beta_2 . \qquad (3.14)$$

A difficulty with this apparently straightforward solution is that the inversion of (3.9) or (3.10) could lead to a function which is not a probability density, or it could yield a probability density but not one which is a probabilistic mixture of three exponentials. In fact, neither of these possibilities is the case, as will be shown by establishing that p_2 and p_3 are positive and subject to the condition $p_2+p_3<1$, and hence can represent probabilities.

Explicit expressions for p_2 and p_3 can be obtained from (3.13) and (3.14) and are given at (3.3) and (3.4). From now on it will be assumed, in accordance with the theorem, that b_2 is the larger of b_2 and b_3, these being obtained by solving the quadratic pair (3.11) and (3.12). To establish that $p_2+p_3<1$, we have, by adding (3.3) and (3.4),

$$p_2+p_3 = \frac{(\alpha_1\beta_1+\alpha_2\beta_2) - (\alpha_1+\alpha_2)\beta_1\beta_2}{(1-b_2)(1-b_3)} . \qquad (3.15)$$

Multiplying out $(1-b_2)(1-b_3)$ in the denominator and using (3.11) and (3.12)

gives, after some rearrangement,

$$p_2+p_3 = 1 - \frac{(1-\beta_1)(1-\beta_2)}{(1-\beta_1)(1-\beta_2) + \alpha_1\beta_1(1-\beta_2) + \alpha_2\beta_2(1-\beta_1)} . \qquad (3.16)$$

The algebraic expression here is clearly positive and less than one, from which it follows that $p_1+p_2<1$.

The positivity of p_2 and p_3 will now be proved by showing that the numerators and denominators of (3.3) and (3.4) are positive. For the denominators, this requires that $0<b_2,b_3<1$ which will be verified by showing that $0<b_2b_3<1$ and $0<(1-b_2)(1-b_3)<1$. The first of these latter two inequalities is obvious from (3.12); for the second consider the expressions

$$\begin{aligned}(1-b_2)(1-b_3) &= 1 - (b_2+b_3) + b_2b_3 \\ &= (\alpha_1\beta_1+1-\beta_1)(\alpha_2\beta_2+1-\beta_2) - (\alpha_1\beta_1)(\alpha_2\beta_2) \qquad (3.17)\end{aligned}$$

after using (3.11) and (3.12), and then

$$\begin{aligned}1 - (1-b_2)(1-b_3) &= b_2 + b_3 - b_2b_3 \\ &= (1-\alpha_1)\beta_1(1-\beta_2) + (1-\alpha_2)\beta_2(1-\beta_1) + \beta_1\beta_2 . \qquad (3.18)\end{aligned}$$

The right hand sides of both (3.17) and (3.18) are obviously positive. This concludes the proof that $0<b_2,b_3<1$ and hence that the denominators of p_2 and p_3 are positive.

For the numerators of p_2 and p_3 to be positive (3.3) and (3.4) indicate that $b = (\alpha_1+\alpha_2)\beta_1\beta_2/(\alpha_1\beta_1+\alpha_2\beta_2)$ must satisfy the inequalities

$$b_3 < b < b_2. \qquad (3.19)$$

At this last stage, explicit expressions for b_2 and b_3 must be used, and from (3.11) and (3.12) are given, after writing

$$s = (1-\alpha_1)\beta_1 + (1-\alpha_2)\beta_2 \quad \text{and} \quad r = (1-\alpha_1-\alpha_2)\beta_1\beta_2 ,$$

by (3.20)

$$b_2 = \{s + (s^2-4r)^{1/2}\}/2 \quad \text{and} \quad b_3 = \{s - (s^2-4r)^{1/2}\}/2 .$$

Then (3.19) is equivalent to

$$-(s^2-4r)^{1/2} < s - 2b < (s^2-4r)^{1/2}$$

or

$$s^2-4r > (s-2b)^2$$

or

$$sb - b^2 - r > 0 . \qquad (3.21)$$

After some algebraic rearrangement the left hand side of (3.21) becomes

$$\alpha_1\alpha_2\beta_1\beta_2(\beta_1-\beta_2)^2/(\alpha_1\beta_1+\alpha_2\beta_2)^2 \qquad (3.22)$$

which is again clearly strictly positive, as was to be proved.

This concludes the proof that p_2 and p_3 are both positive and subject to $p_2+p_3<1$; hence $1-p_2-p_3$, p_2, and p_3 can all be regarded as probabilities. Thus ϵ_n has a proper probability distribution which can be generated as the $(1-p_2-p_3,p_2,p_3)$ mixture of three exponentials of means 1, b_2 and b_3 respectively; further, both b_2 and b_3 are less than unity and $b_2 \neq b_3$.

In special cases there are valid and simpler results for the distribution of ϵ_n. For instance, when $\beta_1=\beta_2=1$, ϵ_n has a simple exponential distribution of mean $(1-\alpha_1-\alpha_2)$. When $\beta_1=\beta_2\neq1$ the innovation has a mixed exponential distribution of the NEAR(1) form given in (2.1) with $\alpha=\alpha_1+\alpha_2$. When $\beta_2=1$, $p_2+p_3=1$ and ϵ_n is the mixture of two exponentials with means b_2 and b_3; this case is used in some of the calculations of Section 9.

4. OTHER USES OF THE NEAR(2) EXPONENTIAL CONSTRUCTION

The NEAR(2) process was established by showing that (3.2) was a valid innovation distribution for the relation (3.1) to give a process with marginal exponential distributions. The distributional assumptions implied by this result can also be taken out of the time series context in which they were derived and viewed generally as a way to combine a pair of (possibly dependent) unit exponential variables (L_1,L_2) with an independent triple of possibly dependent, unit exponential variables (M_1,M_2,M_3) so as to yield a further unit exponential variable. Specifically, with $(\alpha_1,\alpha_2,\beta_1,\beta_2)$ and (b_2,b_3,p_2,p_3) as previously related by (3.3) - (3.7), the Theorem has established that

$$\begin{cases} \beta_1 L_1 & \text{w.p. } \alpha_1 \\ \beta_2 L_2 & \text{w.p. } \alpha_2 \\ 0 & \text{w.p. } 1-\alpha_1-\alpha_2 \end{cases} + \begin{cases} M_1 & \text{w.p. } 1-p_2-p_3 \\ b_2 M_2 & \text{w.p. } p_2 \\ b_3 M_3 & \text{w.p. } p_3 \end{cases} \qquad (4.1)$$

has a unit exponential distribution.

First of all, the idea of "switching" will be illustrated; in the NEAR(2) context, this suggests taking (M_1,M_2,M_3) as $(X_{n-1},X_{n-2},X_{n-3})$ and (L_1,L_2) as (E_n,E_n). Then (4.1) gives the time series model

$$X_n = \begin{cases} X_{n-1} & \text{w.p. } 1-p_2-p_3 \\ b_2 X_{n-2} & \text{w.p. } p_2 \\ b_3 X_{n-3} & \text{w.p. } p_3 \end{cases} + \begin{cases} \beta_1 E_n & \text{w.p. } \alpha_1 \\ \beta_2 E_n & \text{w.p. } \alpha_2 \\ 0 & \text{w.p. } 1-\alpha_1-\alpha_2 \end{cases} . \qquad (4.2)$$

This is a third order autoregression, actually a case of the EAR(3) model cited in Lawrance and Lewis (1980); note, however, that this third order autoregressive exponential process allows zero innovations. Another, better behaved higher order exponential model - in fact a p-th order model - is obtained by the following application of the result (4.1) in its original form (3.1). Let the indices $1,2,\ldots,p$ be partitioned into two non-empty sets I_1 and I_2 of size t_1 and t_2 respectively. Then in the model

$$X_n = \begin{cases} \beta_1' X_{n-1} & \text{w.p. } \alpha_1' \\ \vdots\vdots & \vdots\vdots \\ \beta_p' X_{n-p} & \text{w.p. } \alpha_p' \\ 0 & \text{w.p. } 1-\alpha_1'\ldots-\alpha_p' \end{cases} + E_n \qquad n=0,\pm1,\pm2,\ldots, \qquad (4.3)$$

let $\beta_i'=\beta_1$, $i\epsilon I_1$; $\beta_i'=\beta_2$, $i\epsilon I_2$; $\sum_{i\epsilon I_1}\alpha_i'=\alpha_1$ and $\sum_{i\epsilon I_2}\alpha_i'=\alpha_2$. Then if $\alpha_1+\alpha_2<1$, $0<\beta_1,\beta_2<1$, the distribution of E_n is given by the Theorem. Thus we have a pth order exponential autoregressive process with four parameters. However, while this may seem satisfying it is not clear that four parameters would be sufficient to characterize the sample path behaviour of an exponential process with very high order dependence.

Another use of (4.1) is to allow L_1 and L_2 to both be X_{n-1}, and so obtain a four parameter first order model of the form

$$X_n = \begin{cases}\beta_1 X_{n-1} & \text{w.p. } \alpha_1 \\ \beta_2 X_{n-1} & \text{w.p. } \alpha_2 \\ 0 & \text{w.p. } 1-\alpha_2-\alpha_2\end{cases} + \begin{cases}E_n & \text{w.p. } 1-p_2-p_3 \\ b_2E_n & \text{w.p. } p_2 \\ b_3E_n & \text{w.p. } p_3\end{cases} . \qquad (4.4)$$

Four parameters may seem excessive for a first order autoregressive process but simulations show a wide range of behaviour in sample paths with different choices of parameters. Equation (4.4) in turn suggests a first order model allowing negative dependence. This is obtained by replacing the variable X_{n-1} in (4.4) which is multiplied by β_2 by the antithetic transformation of X_{n-1}, which is $\log\{1-\exp(-X_{n-1})\}$. Two parameter versions of these two first order models could be obtained, for example, by taking $\alpha_1=\alpha_2$, $\beta_1=\beta_2$.

A third type of use of the construction is to give mixed autoregressive moving average models; for this, (L_1,L_2) is (X_{n-1},X_{n-2}) as in (3.1), but (M_1,M_2,M_3) are chosen to be (E_n,E_{n+1},E_{n+1}) for a second order moving average component, or as (E_n,E_{n+1},E_{n+2}) for a third order moving average component; these forward running indices of the innovation sequence are necessary for the required independence in the construction.

Another use of this construction is to obtain an explicit *mixed first order* autoregressive moving average exponential process, which could be contrasted with the *implicit* model given in Jacobs and Lewis (1977). Thus in the basic structure (4.1) replace L_1 by X_{n-1}, L_2 by E_{n-1} and M_1,M_2,M_3 each by E_n; although X_{n-1} and E_{n-1} are dependent, they are independent of E_n as required by the construction.

Out of the time series context, the construction suggests ways to obtain multivariate exponential distributions, rather as in Lawrance and Lewis (1983).

Further possibilities are numerous, but it is not the intention here to exhaustively list them, or to derive the details of those cited. Analysis in the following sections will deal with the basic NEAR(2) model.

5. AUTOCORRELATION STRUCTURE OF THE NEAR(2) PROCESS

In this section we show that the autocorrelations $\rho(\ell)=\text{corr}(X_n,X_{n-\ell})$, $\ell=0,\pm1,\pm2,\ldots$ of the NEAR(2) process satisfy AR(2) Yule-Walker type difference equations; thus the second order dependency of the process is indistinguishable from that of the standard autoregressive model, AR(2). To show this, it is convenient to write the equation (3.1) as a random coefficient additive combination of X_{n-1}, X_{n-2} and E_n. Thus we have the NEAR(2) process in its random coefficient, linear, additive form as

$$X_n = \beta_1 K_n' X_{n-1} + \beta_2 K_n'' X_{n-2} + L_n E_n \qquad n=0,\pm1,\pm2,\ldots, \qquad (5.1)$$

where

$$L_n = \begin{cases} 1 & \text{w.p. } 1-p_2-p_3 \\ b_2 & \text{w.p. } p_2 \\ b_3 & \text{w.p. } p_3 \end{cases}, \qquad n=0,\pm1,\pm2,\ldots, \qquad (5.2)$$

$$\{K_n', K_n''\} = \begin{cases} (1,0) & \text{w.p. } \alpha_1 \\ (0,1) & \text{w.p. } \alpha_2 \\ (0,0) & \text{w.p. } 1-\alpha_1-\alpha_2 \end{cases}; \qquad n=0,\pm1,\pm2,\ldots \qquad (5.3)$$

the i.i.d. sequences $\{L_n\}$ and $\{K_n', K_n''\}$ are assumed to be mutually independent and independent of the independent exponential sequence $\{E_n\}$; the E_n's are assumed to have unit mean, as then do the X_n's by construction.

Now $E(K_n')=\alpha_1$ and $E(K_n'')=\alpha_2$, so that $E(L_n)=1-\alpha_1\beta_1-\alpha_2\beta_2$. Then multiplying X_n in (5.1) by $X_{n-\ell}$ we have, for $\ell \geq 1$,

$$E(X_nX_{n-\ell}) = \alpha_1\beta_1 E(X_{n-1}X_{n-\ell}) + \alpha_2\beta_2 E(X_{n-2}X_{n-\ell}) + E(L_n)E(E_n)E(X_{n-\ell})$$

$$= \alpha_1\beta_1 E(X_{n-1}X_{n-\ell}) + \alpha_2\beta_2 E(X_{n-2}X_{n-\ell}) + 1-\alpha_1\beta_1-\alpha_2\beta_2,$$

so that

$$E(X_nX_{n-\ell})-1 = \alpha_1\beta_1\{E(X_{n-1}X_{n-\ell})-1\} + \alpha_2\beta_2\{E(X_{n-2}X_{n-\ell})-1\}$$

and thus

$$\rho(1) = a_1 + a_2\rho(1), \quad \rho(2) = a_1\rho(1) + a_2, \qquad (5.4)$$

$$\rho(\ell) = a_1\rho(\ell-1) + a_2\rho(\ell-2), \qquad \ell=3,4,\ldots,$$

where $a_1=\alpha_1\beta_1$ and $a_2=\alpha_2\beta_2$. The equations (5.4) are the same as Yule-Walker equations for the standard AR(2) process. The conditions for a solution to exist, namely $a_1+a_2<1$, $a_1-a_2>-1$, $a_2>-1$ are clearly satisfied when the conditions on $\alpha_1,\alpha_2,\beta_1,\beta_2$ given in the Theorem of Section 3 hold.

Note from (5.4) the explicit results,

$$\rho(1) = a_1/(1-a_2) \text{ and } \rho(2) = a_1\rho(1) + a_2. \qquad (5.5)$$

and hence, since $0<a_1,a_2<1$, the restriction of the autocorrelations $\rho(\ell)$ to positive values. The possible region of $(\rho(1),\rho(2))$ values is bounded below by $\rho(2)=(\rho(1))^2$ and otherwise bounded by $\rho(1)>0$ and $\rho(2)<1$. Broadening of the model to negative dependency may be achieved using antithetic ideas, or the bivariate scheme given in Gaver and Lewis (1980), but is not pursued here.

Note too that the parameters in (5.4) enter only as products $a_1=\alpha_1\beta_1$ and $a_2=\alpha_2\beta_2$. Thus for small enough α_1 and α_2, values of β_1 and β_2 greater than unity could be allowed, and (5.4) would still have a stable solution. However, the sequence ϵ_n in the defining equation (3.2) may not exist; it has not been determined whether $(\beta_1\leq 1,\beta_2\leq 1)$ is a necessary condition for this existence.

Specifying allowable values of $\rho(1)$ and $\rho(2)$, as may be done in an initial second order analysis of data, leaves two parameters to be specified in the model, say α_1 and α_2, which could produce very different sample path behaviour in the time series. It is important to notice that this specification of $\rho(1)$ and $\rho(2)$ further constrains the range of possible α_1 and α_2 values. Recalling that $\rho(1)$ and $\rho(2)$ fix $a_1=\alpha_1\beta_1$ and $a_2 = \alpha_2\beta_2$, as well as that $\alpha_1+\alpha_2\leq 1$, it is easily shown that we must have

$$a_1 < \alpha_1 \quad \text{and} \quad a_2 < \alpha_2 \qquad (5.6)$$

which implies that $a_1+a_2\leq\alpha_1+\alpha_2\leq 1$. Thus α_1 and α_2 are forced to lie in a triangular subregion of the triangular (α_1,α_2) region which is bounded below by a_2, bounded on the left by a_1, and bounded above by the line $\alpha_1+\alpha_2=1$. These results are useful, and will be employed in an exploratory analysis of the wind velocity data in Section 9.

6. AN ANALYSIS OF A LONG SERIES OF WIND VELOCITY DATA

6.1 Discussion of the Data

Lewis and Hugus (1982) have given an analysis of a set of 43,800 3-hourly wind velocity readings taken by ship PAPA in the Gulf of Alaska over a 15 year period. After suitable detrending to remove 1 year, 6 month, 12 hour and 6 hour cyclic trend components, a first order autoregressive Gamma model (Lewis, 1981) was fitted to the data, the use of this model being suggested by the shape of the (marginal) histogram of the data (Figure 6.1), the autocorrelation function (Table 6.1) and the shape of the log of the normalized periodogram of the data (Figure 6.2). After detrending there is still a slight 6-hour effect (p=21,900) because this cycle varies in intensity over the 15 years; in what follows this will be ignored and the data will be treated as stationary.

It is not the object here to give the above analysis in detail but rather to give an alternative analysis of the data using NEAR models; this involves a preliminary transformation of the data to an exponential marginal distribution. This is suggested firstly by the fact that a Weibull distribution is commonly used by meteorologists for wind velocity data and secondly by the fact that Weibull and Gamma distributions fit the ship PAPA wind velocity data equally well (Lewis and Hugus, 1982); a power transformation of the Weibull then leads immediately to the desired exponential.

This transformation is preferred to the more usual transformation to normality which, as we shall argue, is not appropriate in this case. The data is in fact finely discretized, with zero values being inlcuded. After a power transform to normality the zero values show up as a group of values still at zero, whereas the non-zero values are shifted away from zero to form the normal part of the data distribution. This zero value problem is not critical with the power transform to exponentiality since this

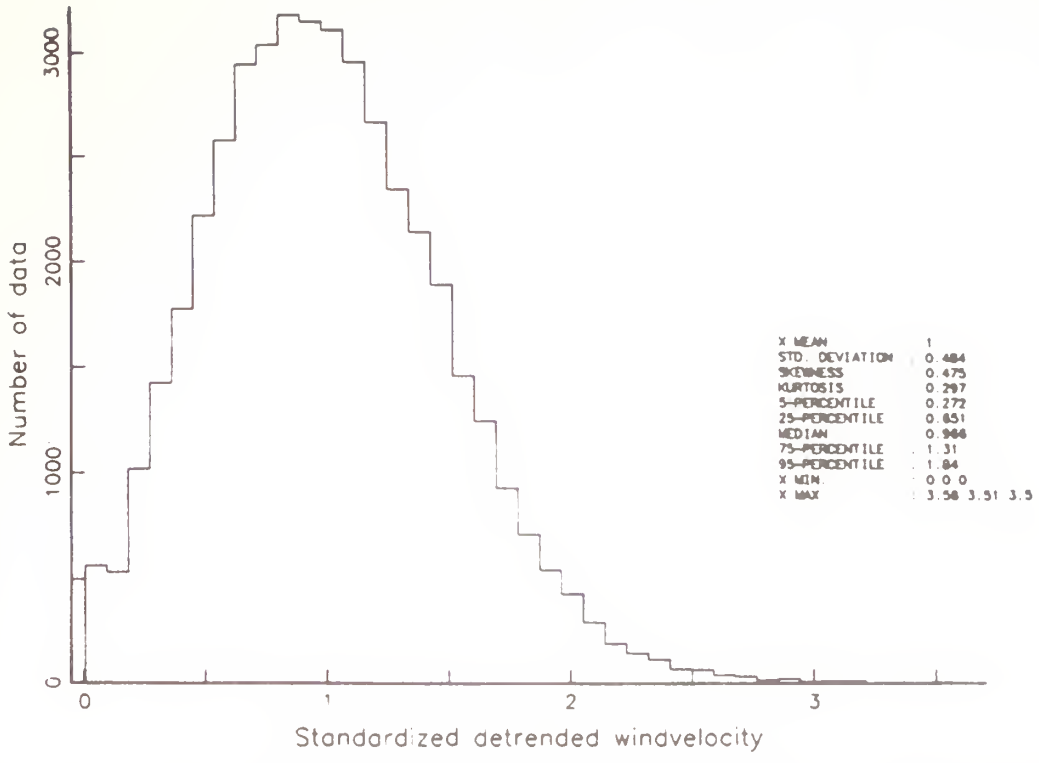

Figure 6.1. Histogram of detrended wind velocity data. Positive skewness is evident, suggesting either a Gamma or Weibull model for the marginal distribution. Sample size N = 43,800.

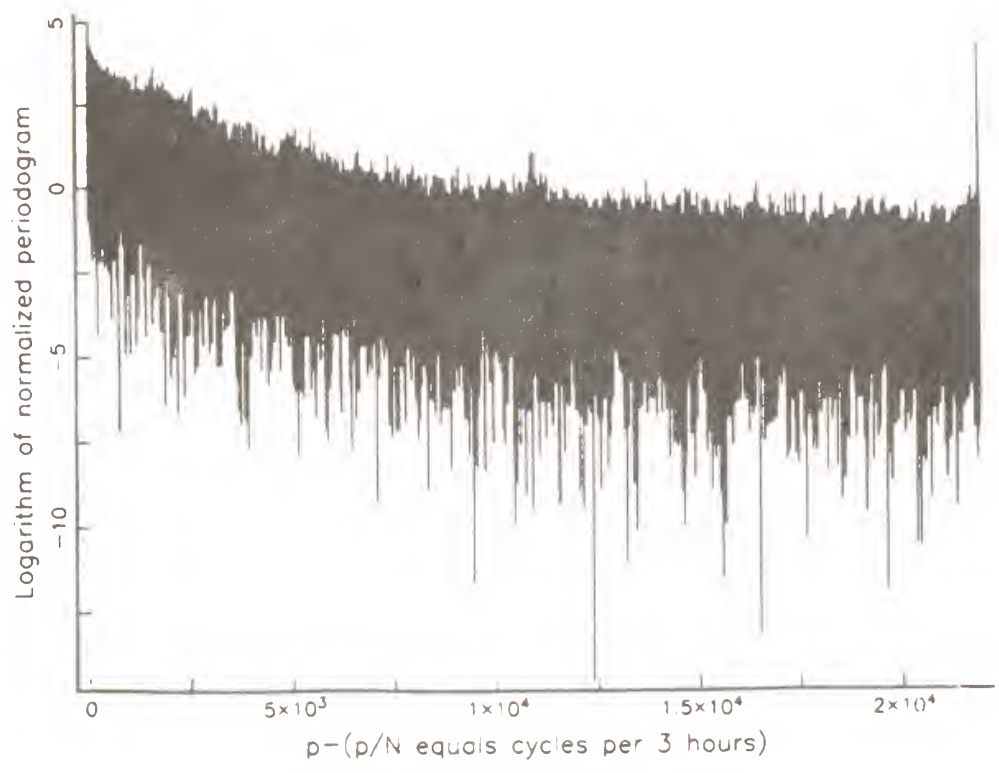

Figure 6.2 Log periodogram of 3-hourly detrended wind velocity data. Note the slight presence of the 6-hours cycle (p=21,900). The periodogram clearly shows that the data is autocorrelated. Sample size N = 43,800.

distribution gives rise to a high proportion of zero values.

It is possible, but extremely tedious, to smooth out the discretization in the data. Even then, however, one can anticipate that after a transformation to marginal normality the time series will be nonlinear; there is no guarantee that such a transformation will produce linearity. Thus we have preferred to transform for marginal exponentiality and attempt to incorporate nonlinearity into the modelling.

The histogram of the transformed data $X_n'=X_n^{2.185}$, is shown in Figure 6.3, where the power transformation to exponentiality has been determined iteratively so that the coefficient of variation of the transformed data is unity. This transformation does affect the correlation structure of the data, as shown in Table 6.1; the table gives comparisons of data and model autocorrelations both before and after the transformation.

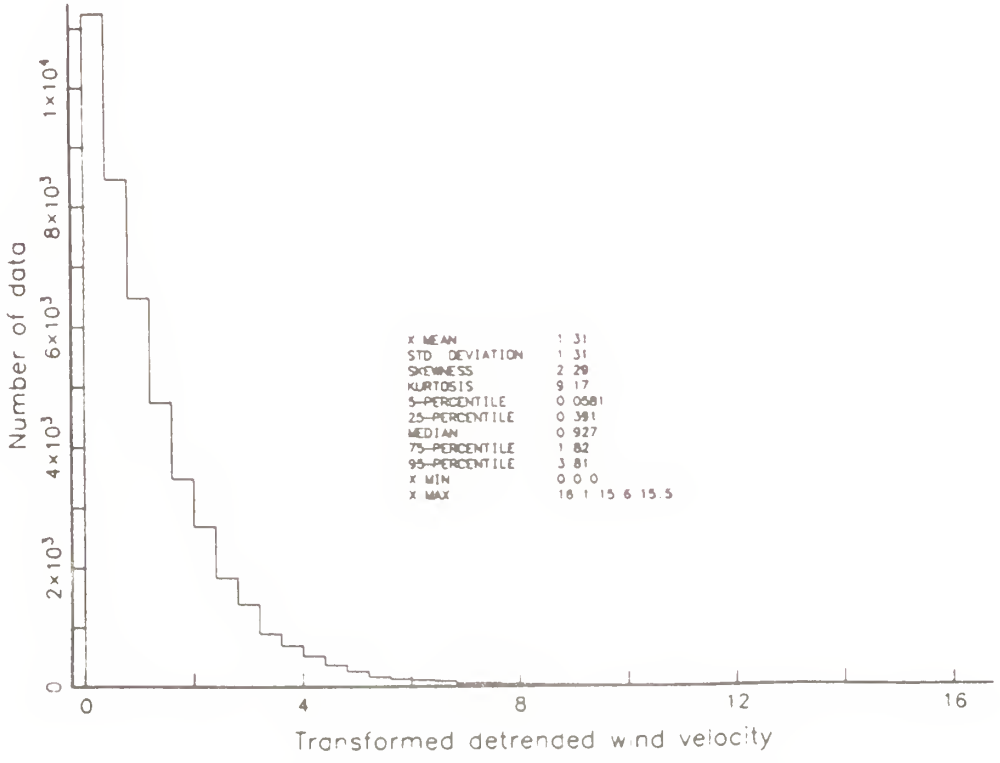

Figure 6.3. Histogram of transformed, detrended wind velocity data. The transformation is $X_n'=X_n^{2.185}$, where the X_n, n=1,2,...,43,800 are the detrended original data. The shape is clearly exponential, but the estimated skewness and kurtosis have values greater than the theoretical values of 2.0 and 6.0, caused by a small percentage of very large values.

Table 6.1

Fit of Wind Speed Data Autocorrelations

Detrended Series $\{X_n\}$				Transformed Detrended Series $\{X_n^{2.185}\}$			
$\rho_1''^{\ell}$	Est.$\rho(\ell)$	$\lvert\Delta\rvert$	$\lvert\Delta\rvert$	$\rho_1\tilde{}^{\ell}$	Est.$\rho(\ell)$	AR(2)	$\lvert\Delta\rvert$
.8194	ρ_1''=.8194	.0000	.0000	.7902	$\rho_1\tilde{}$=.7902	.7902	.0000
.6714	ρ_2''=.6898	.0184	.0346	.6243	$\rho_2\tilde{}$=.6589	.6589	.0000
.5502	ρ_3''=.5635	.0133	.0391	.4933	$\rho_3\tilde{}$=.5324	.5454	.0130
.4508	ρ_4''=.4698	.0190	.0541	.3898	$\rho_4\tilde{}$=.4439	.4519	.0080
.3694	ρ_4''=.3764	.0070	.0431	.3080	$\rho_5\tilde{}$=.3511	.3744	.0233

Table 6.1 The second column shows the estimated autocorrelations for the detrended series; these are close in value to the powers of ρ_1''=0.8194 in the first column, indicating a good fit (Column 3) to a model with AR(1) autocorrelation structure. After transformation, this AR(1) fit is no longer valid (columns 4 and 5) since the absolute values of the differences (Column 4) are now much larger. A much better and adequate fit is obtained with AR(2) autocorrelation structure, as indicated by columns 6, 7 and 8.

Column 2 in Table 6.1 gives the estimated autocorrelations, ρ_ℓ" of the detrended data; the standard error of each of these autocorrelation estimates separately is given approximately by $1/(N)^{1/2}=1/(43,800)^{1/2}$ $=0.005$. The first column of Table 1 gives the fitted autocorrelations for a model with AR(1)-type autocorrelations, just $\rho_1"^{\ell} = (0.8194)^{\ell}$, for lags $\ell=1,2,\ldots,5$. The differences (Column 3) are all very small in practical importance, although some are perhaps statistically significant in view of the large sample size.

Column 6 in Table 6.1 gives the estimated autocorrelations, ρ_ℓ~, for the transformed data; the transformation consistently lowers the autocorrelations. However, columns 5 and 7, which give the fitted AR(1) and AR(2) correlation values, respectively, show that a model with the AR(2) correlation structure is definitely preferable. The fit is borne out by a periodogram plot (not given), and the analysis will be continued on this basis.

Thus, a NEAR(2) model is a candidate for representing the transformed data, and if $\rho(1)$ and $\rho(2)$ are fixed at the estimated values of ρ_1~ = 0.7902 and ρ_2~ = 0.6589, then the corresponding $a_1=\alpha_1\beta_1$ and $a_2=\alpha_2\beta_2$ from (5.4) are, respectively, 0.7175 and 0.0920. There are still two degrees of freedom left in fitting the model, represented by choice of parameters α_1 and α_2 greater than or equal to 0.7175 and 0.0920 respectively, with $\alpha_1+\alpha_2\leq 1$.

Figure 6.4 shows the cumulative periodogram of the usual AR(2) model linear residuals, $R_n=X_n'-a_1X_{n-1}'-a_2X_{n-2}'$, of the transformed data. This is almost straight (ignoring the slight effect at period 6 hours). At this point it might be thought that the usual second order autoregressive model is adequate. We shall however now develop an extended residual analysis for higher order dependence which explores further fitting of the NEAR(2) model to the transformed data.

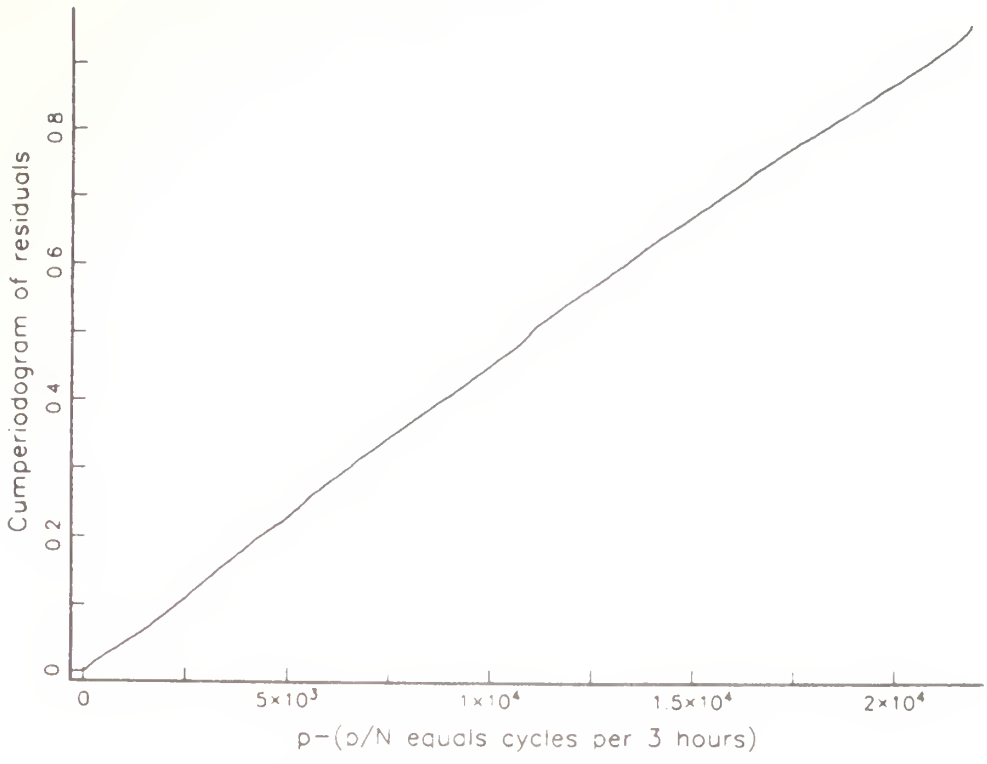

Figure 6.4. Cumulative periodogram of linear residuals R_n, where $R_n=X_n'-0.7175X_{n-1}'-0.0920X_{n-2}'$, for the detrended and transformed wind velocity data. Since the cumulative periodogram is almost straight, the indication is that the residuals are uncorrelated. Note that there is still the suggestion of a six hour effect, indicated by the final upward turn of the graph.

7. A RESIDUAL ANALYSIS FOR THE NEAR(2) MODEL

7.1 General Results

It has already been remarked that the autocorrelations $\rho(\ell)$ are insufficient to describe the dependency structure of NEAR(2) models. A natural next step might be to examine higher order joint moments and their associated spectra. The functions so obtained, such as the bispectrum, (Tukey (1959)), are often found to be difficult to calculate and hard to interpret. Rather than follow this course, a residual analysis for nonlinear autoregressive models is proposed. The thrust of this analysis is that the standard process of fitting and validating a linear autoregressive model should be carried out *beyond* the customary final stage at which uncorrelated residuals are obtained (as in the previous section). The usual presumption is that the residuals are not only uncorrelated but also independent. This need not be the case, as will be exemplified for the wind velocity series. Also, uncorrelated but dependent residuals are obtained for NEAR(2) processes. Thus residuals should be analysed for further dependency. Any found is then evidence that a standard linear, constant coefficient second order autoregressive model is deficient. With normally distributed time series data this might suggest that Gaussian nonlinear modelling should be explored. With data marginally distributed in some other identifiable manner, the exploration of a selected type of nonlinear model with specified marginal distribution and autocorrelation function is suggested; it should then have some higher order dependency properties of its autoregressive residuals in agreement with those of the data. This is the course exemplified here. The proposed residual analysis is further explored in Lawrance and Lewis (1984a).

In Section 9 higher order dependency properties of the uncorrelated residuals from the wind velocity data are compared with those derived in Section 8 for the NEAR(2) model. This stage can be informative from both

exploratory and estimation considerations, and can be thought of as part of the model-refinement process common to much statistical methodology.

It might be thought that the class of NEAR(2) models could be subjected to a residual analysis in the standard manner or using more general forms of residuals that have been studied by Cox and Snell (1968). However, considering the NEAR(2) model in its random coefficient form (5.1), the independent innovation is now trivariate, consisting of (K_n, K_n', L_nE_n). No way of estimating this trivariate distribution based on realized $\{X_n\}$ has been found, even assuming knowledge or estimation of the model parameters. The linear autoregressive residuals are never-the-less available, and given as in the previous section by

$$R_n=X_n-a_1X_{n-1}-a_2X_{n-2},\ a_1=\alpha_1\beta_1,\ a_2=\alpha_2\beta_2,\ n = 0,\pm1,\pm2,... \qquad (7.1)$$

We now show that these residuals are uncorrelated for the NEAR(2) process.

7.2 **The Residual Theorem**

Theorem: The residuals $(R_n,R_{n+\ell})$, given by (7.1), are uncorrelated for $\ell=\pm1,\pm2,...$.

Proof: The autocovariances of the residuals (7.1) may, for $\ell\geq1$, be written,

$$\begin{aligned}\mathrm{Cov}(R_n,\ R_{n+\ell}) &= \mathrm{Cov}(X_n,\ R_{n+\ell})-a_1\mathrm{Cov}(X_{n-1},\ R_{n+\ell})-a_2\mathrm{Cov}(X_{n-2},\ R_{n+\ell}) \qquad (7.2)\\ &= \mathrm{Cov}(X_n,\ R_{n+\ell})-a_1\mathrm{Cov}(X_n,\ R_{n+\ell+1})-a_2\mathrm{Cov}(X_n,\ R_{n+\ell+2}),\end{aligned}$$

since the $\{X_n\}$ process and consequently the $\{R_n\}$ process is stationary. The covariances on the right hand side are all of the same type and given by

$$\begin{aligned}\mathrm{Cov}(X_n,\ R_{n+\ell}) &= \mathrm{Cov}\{X_n,\ (X_{n+\ell} - a_1X_{n+\ell-1} - a_2X_{n+\ell-2})\}\\ &= \{\mathrm{Var}(X)\}(\rho(\ell) - a_1\rho(\ell-1) - a_2\rho(\ell-2)),\quad \ell=1,2,... \qquad (7.3)\end{aligned}$$

An identical result also holds for ℓ's less than zero. Hence by the Yule-Walker equations (5.4), the expression in brackets is zero, and so

$$\mathrm{Corr}(R_n,\ R_{n+\ell}) = 0, \qquad \ell=\pm1,\pm2,..., \qquad (7.4)$$

as was to be proved. That these residuals are uncorrelated is an immediate consequence of the autocorrelations following Yule-Walker equations; this

emphasizes that residuals of the form (7.1) will be uncorrelated for *any* model whose autocorrelations satisfy Yule-Walker equations. Equivalently, as was the case for the transformed detrended wind velocity data (Figure 6.4), the cumulative periodogram of the residuals will plot linearly.

Dependency analysis of the uncorrelated residuals $\{R_n\}$, n=1,2,... could begin with scatter plots of the low-lag adjacent values; any patterns or concentrations are suggestive of dependency. Several further methods for the detection of dependency could be proposed, but the only one pursued here involves the squares of residuals; it is then applied in Section 9 to a continued analysis of the wind velocity data.

8. CROSSCOVARIANCE ANALYSIS OF $\{R_n\}$ AND $\{R_n^2\}$ FOR THE NEAR(2) MODEL

After the satisfactory fit to data of an ordinary linear model, the residuals, R_n, should be independent; this is conveniently investigated by seeking straight cumulative periodograms for the residuals and for the squared residuals; for the wind velocity data, Figure 8.1 shows that the squared residuals have an obviously curved cumulative periodogram. Thus a linear AR(2) model for this data is definitely not adequate. As a method for probing model validity, the examination of squared residuals has previously been employed by McLeod and Li (1983), following Granger and Andersen (1978); these latter authors investigated bilinear modelling of uncorrelated but dependent residuals from ARMA models, with a view to improved forecasting.

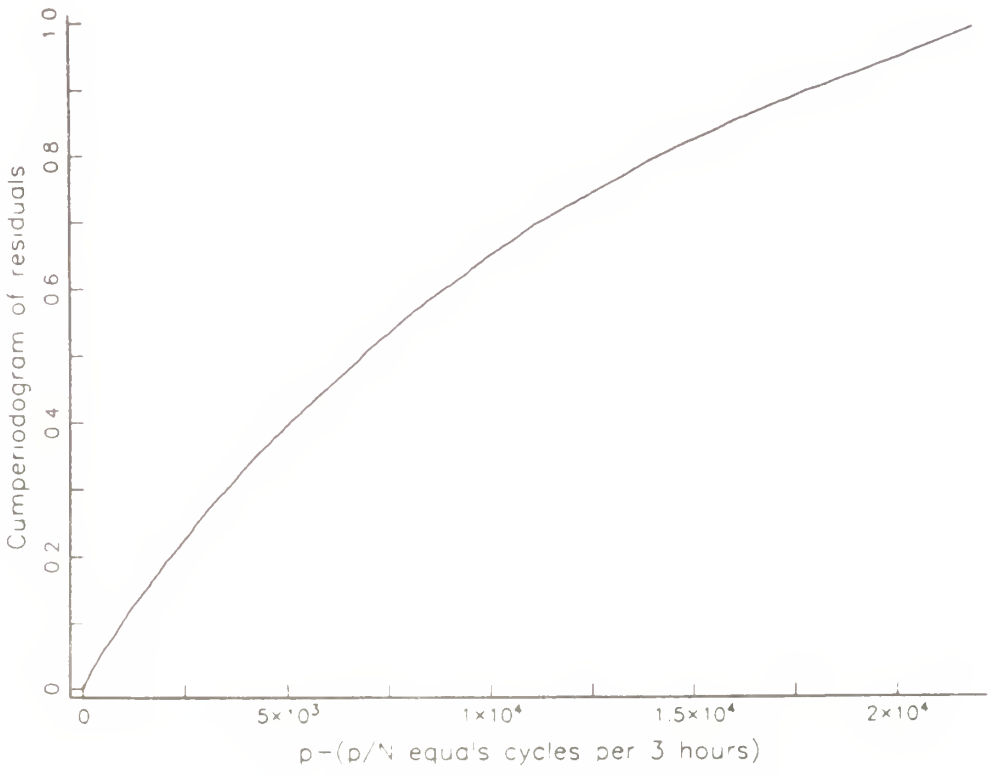

Figure 8.1. Cumulated periodogram values for the squared residuals, R_n^2, of the transformed detrended wind velocity data. The underlying spectrum is clearly not flat; the Kolmogorov-Smirnov test statistic is 6.14 which is very high compared to the upper .99 quantile of the statistic which is 1.618.

Whilst autocorrelations of the squared residuals can be considered, for the NEAR(2) model this involves computation of 36 terms, mostly distinct types of 4th order moments. A more tractable suggestion which involves only third order moments, and which is thus the next step up after ordinary autocorrelations, is to use the *crosscovariances* of the $\{R_n\}$ and $\{R_n^2\}$ sequences; apart from lag 0, zero values would be found for linear models. We use crosscovariances rather than crosscorrelations of $\{R_n\}$ and $\{R_n^2\}$ because crosscorrelations need $Var(R_n^2)$ which involves 4th order moments of the X_n process.

The starting point for the calculation of $Cov(R_n^2, R_{n-\ell})$ is to note from the definition of R_n at (7.1) that

$$E(R_n^2 R_{n-\ell}) = E(R_n^2 X_{n-\ell}) - a_1 E(R_n^2 X_{n-\ell-1}) - a_2 E(R_n^2 X_{n-\ell-2}), \quad \ell=0,\pm1,\pm2,... \quad (8.1)$$

whence there is the structural form,

$$Cov(R_n^2, R_{n-\ell}) = Cov(R_n^2, X_{n-\ell}) - a_1 Cov(R_n^2, X_{n-\ell-1}) - a_2 Cov(R_n^2, X_{n-\ell-2}). \quad (8.2)$$

Calculation of the covariance terms in (8.2) requires the expanding out of R_n^2, taking expectations and expression in terms of covariances. Thus

$$\begin{aligned} R_n^2 X_{n-\ell} = {} & X_n^2 X_{n-\ell} + a_1^2 X_{n-1}^2 X_{n-\ell} + a_2^2 X_{n-2}^2 X_{n-\ell} \\ & - 2a_1 X_n X_{n-1} X_{n-\ell} - 2a_2 X_n X_{n-2} X_{n-\ell} \\ & + 2a_1 a_2 X_{n-1} X_{n-2} X_{n-\ell}, \qquad \ell=0,\pm1,\pm2,... \quad (8.3) \end{aligned}$$

The conversion to covariance yields

$$\begin{aligned} Cov(R_n^2, X_{n-\ell}) = {} & J_1(\ell) + a_1^2 J_1(\ell-1) + a_2^2 J_1(\ell-2) \\ & - 2a_1 J_2(\ell) - 2a_2 J_3(\ell) + 2a_1 a_2 J_2(\ell-1), \quad \ell=0,\pm1,\pm2,... \quad (8.4) \end{aligned}$$

where

$$J_1(\ell) = Cov(X_n^2, X_{n-\ell}); \quad J_2(\ell) = Cov(X_n X_{n-1}, X_{n-\ell}); \quad J_3(\ell) = Cov(X_n X_{n-2}, X_{n-\ell}). \quad (8.5)$$

We thus see the types of third order joint moments for the $\{X_n\}$ process which are involved in the $Cov(R_n^2, R_{n-\ell})$ calculation; these are general results.

For the NEAR(2) model each of these joint moments has to be obtained using a difference equation. Considering for illustration $J_1(\ell)$ for

positive lags, square each side of the NEAR(2) defining equation (2.2) and multiply by $X_{n-\ell}$. After converting to the required covariances, the recursion is found to be

$$J_1(\ell) = a_1\beta_1 J_1(\ell-1) + a_2\beta_2 J_1(\ell-2) + 2(1-a_1-a_2)\rho(\ell), \qquad \ell=1,2,... \qquad (8.6)$$

This equation is given for illustration: there are similar equations for $J_2(\ell)$ and $J_3(\ell)$, and various special cases. The complete algorithm for computing $Cov(R_n^2, R_{n-\ell})$ for $\ell \geq 1$ for the NEAR(2) process is given in the Appendix.

The other half of the crosscovariance function of $\{R_n\}$ and $\{R_n^2\}$ is $Cov(R_n^2, R_{n-\ell})$ for $\ell \leq 1$, or equivalently $Cov(R_n^2, R_{n+\ell})$ for $\ell \geq 1$; we now show that this is zero. The key result in establishing this fact is obtained by first defining $G_{n-\ell}$ as either $X_{n-\ell}^2$, $X_{n-\ell}X_{n-\ell-1}$, or $X_{n-\ell}X_{n-\ell-2}$, for $\ell \geq 1$ and showing that $Cov(R_n, G_{n-\ell})$ for $\ell \geq 1$ are all zero.

First note that

$$\begin{aligned} Cov(R_n, G_{n-\ell}) &= E[\{(X_n - a_1X_{n-1} - a_2X_{n-2}) - E(R_n)\}\{G_{n-\ell} - E(G_{n-\ell})\}] \\ &= E[\{X_n - a_1X_{n-1} - a_2X_{n-2}\}\{G_{n-\ell} - E(G_{n-\ell})\}]. \end{aligned} \qquad (8.7)$$

Now substitute for X_n from (5.1) to obtain

$$Cov(R_n, G_{n-\ell}) = E[\{(\beta_1K_n' - a_1)X_{n-1} + (\beta_2K_n'' - a_2)X_{n-2} + L_nE_n\}\{G_{n-\ell} - E(G_{n-\ell})\}]. \qquad (8.8)$$

Since (K_n', K_n'') are independent of (X_{n-1}, X_{n-2}) and L_nE_n is independent of $G_{n-\ell}$, the right hand side of (8.8) may be written

$$Cov(R_n, G_{n-\ell}) = E(\beta_1K_n' - a_1)Cov(X_{n-1}, G_{n-\ell}) + E(\beta_2K_n'' - a_2)Cov(X_{n-2}, G_{n-\ell}). \qquad (8.9)$$

By the definition of (K_n', K_n'') at (5.3), $E(\beta_1K_n' - a_1)$ and $E(\beta_2K_n'' - a_2)$ are both zero, and hence

$$Cov(R_n, G_{n-\ell}) = 0, \qquad \ell=1,2,... \qquad (8.10)$$

Finally, we note that

$$\begin{aligned} Cov(R_n^2, R_{n+\ell}) &= Cov(R_n, R_{n-\ell}^2) \\ &= Cov(R_n, X_{n-\ell}^2) + a_1^2Cov(R_n, X_{n-\ell-1}^2) + a_2^2Cov(R_n, X_{n-\ell-1}^2) \\ &\quad - 2a_1Cov(R_n, X_{n-\ell}X_{n-\ell-1}) - 2a_2Cov(R_n, X_{n-\ell}X_{n-\ell-2}) \\ &\quad + 2a_1a_2Cov(R_n, X_{n-\ell-1}X_{n-\ell-2}). \qquad \ell=1,2,... \end{aligned} \qquad (8.11)$$

By (8.10) all the covariances in (8.11) are zero, and the desired result is established.

In a similar way it may be seen that $Cov(R_n^r, R_{n+\ell})$ for $\ell \geq 1$ and all positive integers r are also zero; this does not, however, imply that R_n and $R_{n+\ell}$ are independent. All joint moments would have to be zero but in particular this is not the case for $Cov(R_n^2, R_{n-\ell})$, $\ell \geq 1$, as results in the Appendix indicate.

It should perhaps also be noted explicitly that in using the residuals $\{R_n\}$ that the coefficients a_1 and a_2 will have been estimated; this may in fact induce some correlation between R_n^2 and $R_{n+\ell}$, but with long series of data the effect should be very small.

9. FURTHER ANALYSIS AND MODELLING OF THE WIND VELOCITY DATA

Dependence in the uncorrelated second order residuals of the transformed detrended wind velocity data has already been demonstrated by Figure 8.1. Further evidence of this is provided by the non-zero crosscovariances of $(R_n, R_{n+\ell^2})$ given in Figure 9.1. The corresponding theoretical crosscovariances for the NEAR(2) model will next be presented, having been computed using the algorithm described in Section 8 and the Appendix.

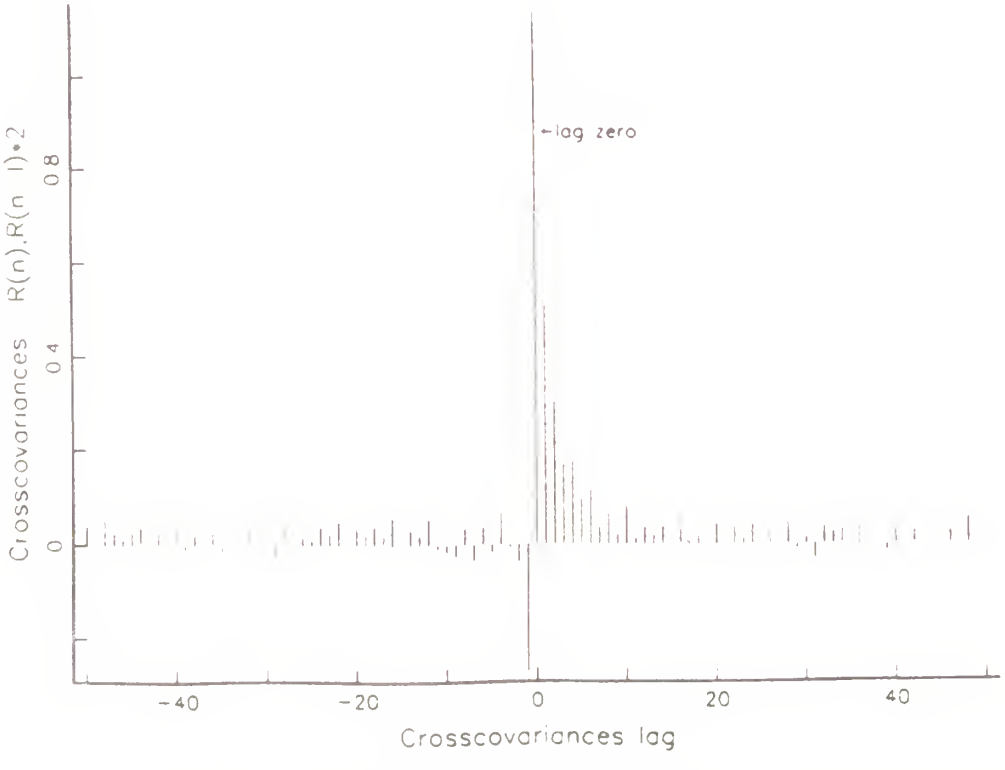

Figure 9.1. Crosscovariances between the second order residuals, R_n, and the squared second order residuals, $R_{n+\ell^2}$, for the transformed detrended wind velocity data. Sample size is N=43,800. There is a strong negative value at lag minus one and pronounced positive values at the first few positive lags.

At this point, it will be recalled, the NEAR(2) model has not been fitted in terms of all 4 parameters; the residuals involve the model parameters only through $a_1 = \alpha_1\beta_1$ and $a_2 = \alpha_2\beta_2$ and $\alpha_1,\beta_1,\alpha_2,\beta_2$ have not been separately estimated. In the present rather exploratory analysis the estimation problem will be circumvented; the crosscovariances of $(R_n, R_{n+\ell^2})$ will be given for four representative sets of parameter values in the reduced allowable region, as constrained by $(\alpha_1 > a_1, \alpha_2 > a_2, \alpha_1+\alpha_2 < 1)$, according to (5.6). For the transformed detrended wind velocity data, a_1=0.7175 and a_2=0.0920, and the four chosen sets of $\alpha_1,\alpha_2,\beta_1,\beta_2$ together with their associated (p_2,p_3) and (b_2,b_3) sets are given in Table 9.1.

Table 9.1

Parameters Sets for NEAR(2) Calculations

	α_1	α_2	β_1	β_2	p_2	p_3	b_2	b_3
A	.72	.0920	.9965	1.0000	.0038	.9962	.9996	.1874
B	.76	.0920	.9441	1.0000	.0585	.9415	.9940	.1406
C	.77	.1400	.9318	0.6571	.0542	.8610	.7008	.0786
D	.88	.0920	.8153	1.0000	.1744	.8256	.9826	.0232

In Figure 9.2 the crosscovariances of $(R_n, R_{n+\ell^2})$ for each of these four cases are presented; the NEAR(2) process has been taken with mean of 1.31 so as to correspond to the transformed detrended wind velocity data. First it will be recalled that for the usual linear AR(2) process, these crosscovariances would all be zero, apart from that one at lag zero which is giving a non-standard measure of skewness of the residuals; Case D is nearest to this situation, although the model cannot reduce to the linear AR(2) model; it can reduce to the linear AR(1) model and case D is also nearest to this situation.

Case C is nearest in qualitative behaviour to the crosscovariances of the wind velocity data given in Figure 9.1, but as we have seen, cannot exhibit the negative crosscovariance shown at lag minus one. This is an

important deficiency of modelling by the NEAR(2) process although matters would have been considerably worse for data with a double sided $Cov(R_n^2, R_{n+\ell})$ function. The other cases (A and B) illustrate some of the diversity of behaviour producible by the NEAR(2) model; more investigation of these aspects would be valuable. Further analysis would require formal estimation of all four parameters, or perhaps formal estimation of (α_1,α_2) after the fixing of a_1 and a_2 at their values determined by estimates of $\rho(1)$ and $\rho(2)$.

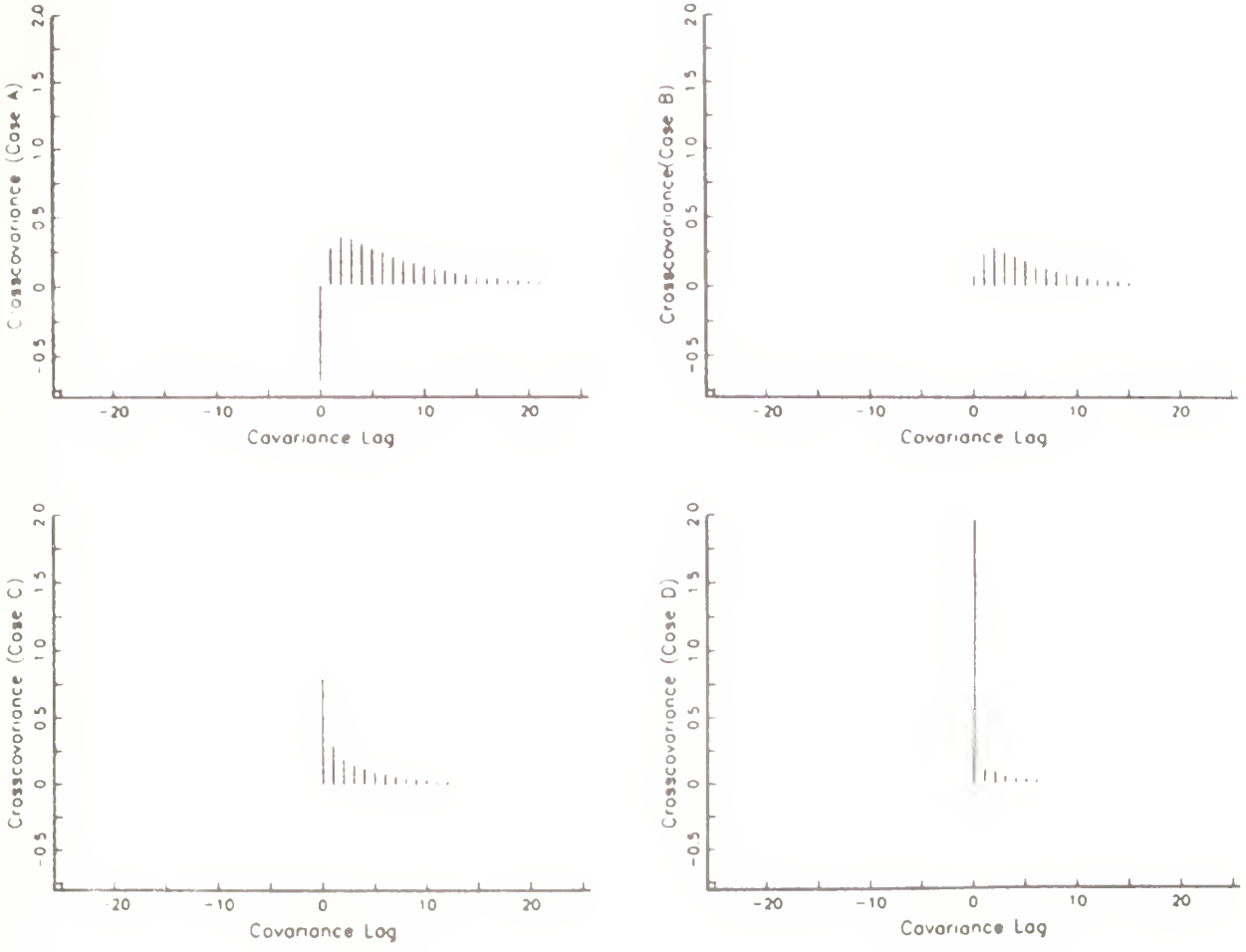

Figure 9.2. Crosscovariances, for the NEAR(2) model, between the second order residuals, R_n, and second order residuals $R_{n+\ell}^2$, with a_1=0.7175 and a_2=0.0920. The sets of parameter values $(\alpha_1,\alpha_2,\beta_1,\beta_2,p_2,p_3,b_2,b_3)$ are described in Table 9.1; case C is closest to that of the wind velocity data given in Figure 9.1. Case D is nearest to the linear EAR(1) model. The other two cases illustrate more of the diversity of behaviour which can be shown by this NEAR(2) crosscovariance function. The lag zero covariance is a measure (not the usual one) of skewness of the residuals.

10. CONCLUSIONS AND FURTHER ANALYSIS

The very broad four parameter NEAR(2) time series model having exponential marginals and the correlation structure of a linear AR(2) model has been established. A preliminary fit of the NEAR(2) model has been made to a very long series of wind speed data, the data having been detrended and transformed so as to have exponentially distributed marginals. A residual analysis has been based on the crosscovariances between the residuals and squared residuals, and its utility in probing higher order dependence in the $\{X_n\}$ process has been demonstrated; in particular it has highlighted a strong difference in higher order crosscovariance of the data and the NEAR(2) model at lag minus one.

A possible extension of the NEAR(2) model which retains the marginal exponential distribution of the data is obtained by noting that the theorem in Section 3 does not require independence of the random coefficient sequences $\{K_n'\}$ and $\{K_n''\}$. By allowing these to be, say, moving average sequences, the higher order structure of the $\{X_n\}$ process might be extended to accomodate the negative-valued spike in Figure 9.1 at lag minus one. This approach has been taken by McKenzie (1981) but it is not known how tractable the resulting structure of the NEAR(2) model would be.

Finally, we remark that a likelihood conditional on the first two values, X_1 and X_2, can be written down for the process. However the likelihood function is difficult to use because it becomes infinite at parameter values on the border of the parameter space corresponding to an EAR(?) model. These considerations suggest that a reduced three parameter model is needed which avoids these singularities but still allows for a higher order dependency in the model. Such a likelihood analysis would also need model validation which could be based on the higher order residual analysis presented here.

An extension of the residual analysis given here based on reversed

residuals is possible (Lawrance and Lewis, 1984b); however, these were not used because residual analysis has already turned up discrepancies between the data and the model which need further explanation.

Appendix

Algorithm for Computing the Non-Negative half of the {Residual, Residual Squared} Crosscovariance Function for the NEAR(2) Model

Input $\alpha_1, \alpha_2, \beta_1, \beta_2, L$

0. $a_1=\alpha_1\beta_1$; $a_2=\alpha_2\beta_2$;

$\rho(0) = 1$; $\rho(1) = a_1/(1-a_2)$; $\rho(2) = a_1\rho(1) + a_2$

$\rho(\ell) = a_1\rho(\ell-1) + a_2\rho(\ell-2)$, for $\ell=2,3,\ldots,L$.

1. $J_1(0) = 4$; $J_1(1) = [4a_1(\beta_1+a_2\beta_2) + 2(1-a_1-a_2)]/(1-a_2^2\beta_2)$.

2a. $J_1(\ell) = a_1J_1(\ell+1) + a_2J_1(\ell+2)$ for $\ell=-1,-2$.

2b. $J_1(\ell) = a_1\beta_1J_1(\ell-1) + a_2\beta_2J_1(\ell-2) + 2(1-a_1-a_2)\rho(\ell)$ for $\ell=2,3,\ldots,L$.

3. $J_2(0) = J_1(1) - \rho(1) + 1$; $J_2(1) = J_1(-1) - \rho(1) + 1$.

4a. $J_2(-1) = a_1J_2(0) + a_2J_2(1)$.

4b. $J_2(\ell) = a_1J_1(\ell-1) + a_2J_2(\ell-1) + 2a_1 + [1+\rho(1)]a_2$

$-[1+\rho(1)] + [1 + \rho(\ell-1)](1-a_1-a_2)$ for $\ell=2,3,\ldots,L$.

5. $J_3(0) = J_1(2) - \rho(2) + 1$; $J_3(1) = J_2(2)+\rho(1) - \rho(2)$;

$J_3(2) = J_1(-2) - \rho(2) + 1$.

6. $J_3(\ell) = a_1J_2(\ell-1) + a_2J_1(\ell-2) + [1+\rho(1)]a_1 + 2a_2$

$+ [1+\rho(\ell-2)](1-a_1-a_2) - [1+\rho(2)]$ for $\ell=3,4,5,\ldots,L$.

7. $J(\ell) = J_1(\ell) + a_1^2J_1(\ell-1) + a_2^2J_1(\ell-2) - 2a_1J_2(\ell)$

$- 2a_2J_3(\ell) + 2a_1a_2J_2(\ell-1)$ for $\ell=0,1,\ldots,L$.

8. $Cov(R_n^2, R_{n-\ell}) = J(\ell) - a_1J(\ell+1) - a_2J(\ell+2)$ for $\ell=0,1,\ldots,L$.

ACKNOWLEDGEMENTS

The research of P.A.W. Lewis was supported by the Office of Naval Research under Grant NR-042. The graphs were produced by an experimental APL package from IBM which the Naval Postgraduate School is using under an agreement with the IBM Research Centre, Yorktown Heights, N.Y.

REFERENCES

Cox, D.R. and Snell E.J. (1968). A general definition of residuals (with discussion). *J. R. Statist. Soc., B, 30, 248-275*

Gaver, D.P. and Lewis, P.A.W. (1980). First order autoregressive gamma sequences and point processes. *Adv.Appl.Prob., 12,* 724-745.

Granger, C.W. and Andersen, A.P. (1978). *An Introduction to Bilinear Time Series Models.* Vandenhoeck and Ruprecht, Gottingen.

Jacobs, P.A. and Lewis, P.A.W. (1977). A mixed autoregressive-moving average exponential sequence and point process (EARMA(1,1)). *Adv. Appl. Prob., 9, 87 104.*

Jacobs, P.A. and Lewis, P.A.W. (1983). Stationary discrete autoregressive - moving average time series generated by mixtures. *J. Time Series Analysis, 4,* 18-36.

Lawrance, A.J. (1980a). The mixed exponential solution to the first-order autoregressive model. *J. Appl. Prob., 12,* 522-546.

Lawrance, A.J. (1980b). Some autoregressive models for point processes. *Point Processes and Queueing Problems* (Eds. P. Bartfai and J. Tomko), Eds 2'7-275. Colloquia Mathematica Societatis Janos Bolyai 24, North Holland, Amsterdam,

Lawrance, A.J. (1982). The innovation distribution of a gamma distributed autoregressive process. *Scand. J. Statist., 9,* 234-236.

Lawrance, A.J. and Lewis, P.A.W. (1980). The exponential autoregressive moving average EARMA(p,q) process. *J.R. Statist. Soc. B, 42,* 150-161.

Lawrance, A.J. and Lewis, P.A.W. (1981). A new autoregressive time series model in exponential variables (NEAR(1)). *Adv. Appl.Prob., 13*, 826-845.

Lawrance, A.J. and Lewis, P.A.W. (1982). A mixed exponential time series model. *Management Science, 28*, 1045-1053.

Lawrance, A.J. and Lewis, P.A.W. (1983). Simple dependent pairs of exponential and uniform random variables. *Operations Research, 31*, 1179-1197.

Lawrance, A.J. and Lewis P.A.W. (1984a). Higher order residual analysis for nonlinear time series with autoregressive correlation structures. Preprint.

Lawrance, A.J. and Lewis, P.A.W. (1984b). Reversed residuals in time series analysis. In preparation.

Lewis, P.A.W. (1981). Simple multivariate time series for simulations of complex systems. In *Proc.Winter Simulation Conference* (Eds, T.I. Oren, C.M. Delfosse and C.M. Shub), 389-390. IEEE Press: New York.

Lewis, P.A.W. and Hugus, D.K. (1982). An analysis of 15 years of wind velocity data from ship PAPA. In D.K. Hugus's Ph.D. thesis, *Extension of Some Models for Positive Valued Time Series*. Naval Postgraduate School, Monterey.

McKenzie E. (1981). Extending the correlation structure of exponential autoregressive moving average processes. *J. Appl. Prob., 18*, 181-189.

McKenzie, E. (1982). Product autoregression: A time series characterization of the gamma distribution. *J. Appl. Prob., 19*, 463-468.

McLeod, A.I. and Li, W.K. (1982). Diagnostic checking ARMA time series models using squared-residuals autocorrelations. *J. Time Series Analysis, 4*, 269-273.

Nicholls, D.F. and Quinn, B.G. (1982). *Random Coefficient Autoregressive Models: An Introduction*. Springer Verlag, Heidelberg.

Raftery, A. (1981). Generalized non-normal time series models. In *Time*

Series Analysis: Theory and Practice 1, (Ed. O.D. Anderson), 621-640. North Holland, Amsterdam.

Tukey, J.W. (1959) An introduction to the measurement of spectra. *Probability and Statistics* (Ed. U. Grenander), 300-330. Wiley, New York.

DISTRIBUTION LIST

	NO. OF COPIES
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Library Code 0142 Naval Postgraduate School Monterey, CA 93943	2
Research Administration Code 012A Naval Postgraduate School Monterey, CA 93943	1
Library Code 55 Naval Postgraduate School Monterey, CA 93943	1
Professor P.A.W. LEWIS Code 55Lw Naval Postgraduate School Monterey, CA 93943	150

U21980

DUDLEY KNOX LIBRARY - RESEARCH REPORTS
5 6853 01060233 7
U215869